혼합형 이미지 메타데이타를 이용한 지능적 이미지 검색 시스템 설계 및 구현

홍성용[†] · 나연묵[††]

요 약

멀티미디어 데이타의 중요성과 활용도가 증가됨에 따라 데이타베이스에서 멀티미디어 데이타를 표현하고 관리하기 위한 연구가 필요하게 되었다. 따라서, 본 논문은 혼합형 이미지 메타데이타를 이용하여 이미지 데이타 베이스에서 이미지 데이타에 대한 검색 기법을 지능화 하고 시스템을 효율적으로 관리할 수 있는 시스템을 제시한다. 개념 계층을 기반으로 이미지 정보를 지능적인 형태로 저장, 검색하고, 이미지 마이닝을 하기위해 사용자의 정보와 사용자가 이용한 질의 정보 그리고 이미지 정보를 통합하여 검색한다. 즉, 사용자 메타데이타와 이미지 메타데이타를 추출하고 통합하여 사용자에게 좀더 많은 부가 정보를 제공할 수 있는 지능적 이미지 검색 시스템을 제공한다.

A Design and Implementation of Intelligent Image Retrieval System using Hybrid Image Metadata

Sungyong Hong[†] and Yunmook Nah[††]

ABSTRACT

As the importance and utilization of multimedia data increases, it becomes necessary to represent and manage multimedia data within database systems.

In this paper, we designed and implemented an image retrieval system which support efficient management and intelligent retrieval of image data using concept hierarchy and data mining techniques. We stored the image information intelligently in databases using concept hierarchy. To support intelligent retrievals and efficient web services, our system automatically extracts and stores the user information, the user's query information, and the feature data of images. The proposed system integrates user metadata and image metadata to support various retrieval methods on image data.

1. 서 론

최근 몇 년간 멀티미디어에 대한 전 세계의 관심이 집중되면서 멀티미디어 데이타베이스 구축을 위한 멀티미디어 데이타 모델링, 검색, 인터페이스 등이 중요한 연구 분야로 대두되고 있다.

이미지 데이타를 검색하기 위한 가장 기본적인 방법으로는 이미지 정보의 내용을 텍스트로 기술하고 그 기술된 내용에 의해서 검색을 하는 방식이 보편적으로 사용되고 있다. 이미지 데이타는 기존의 텍스트 기반의 데이타에 비하여 대용량적인 특성과 비정형적인 특성을 가지고 있어서 신속하고 효율적인 검색에 많은 어려움이 있다. 최근에 이러한 이미지 데이타를 신속하고 효율적으로 검색하기 위하여 다양한 내용 기반 이미지 검색 기법(content-based image retrival)과 데이타 마이닝 또는 OLAP(on-line analysis processing)과 같은 지식을 기반으로 하는 검색 방법들이 연구되고 있다. 또한, MPEG-7과 같은 멀

본 연구는 한국과학재단의 특정기초연구과제 (과제번호 : 98-0102-06-01-3) 연구비 지원에 의한 것임.

[†] 준회원, 단국대학교 컴퓨터공학과 박사과정
[††] 정회원, 단국대학교 컴퓨터공학과 부교수

티미디어 정보를 기술(description)하기 위한 표준화가 활발하게 연구되어지고 있다[1].

그러나 기존의 이미지 검색 방법들은 과거의 검색 결과에 대한 사용자의 만족도를 고려 하지않고 있으며, 통합적으로 이미지의 특성을 이용하지 못하여 검색 효율이 떨어진다는 단점을 지니고 있다[2,3]. 사용자가 이용하고자 하는 이미지를 빠르고 신속하게 제공하기 위해서는 좀 더 지능적인 검색 기법이 요구된다. 따라서, 사용자가 검색했던 정보나 사용자의 부가정보를 데이타베이스에 저장하고 이러한 정보를 이용하여 기존의 이미지 검색 방법과 통합한다면, 지능적인 이미지 검색 뿐만 아니라 사용자의 만족도를 높일 수 있다. 또한, 웹 상에서 이미지를 찾기 위한 다른 사용자에게 중요한 정보로 제공될 수 있을 뿐만 아니라, 이미지를 검색하려고 하는 사용자는 자신과 유사한 환경이나 취미, 연령을 가진 다른 사용자가 검색했던 이미지를 좀 더 신속하고 효율적으로 검색할 수 있게 함으로써 기존에 이미지 데이타만을 고려하여 검색하던 시스템보다 지능적이고 효율적인 검색을 수행할 수 있다. 지능적 이미지 검색 시스템을 개발하기 위한 기본적인 아이디어는 웹상에서 어느 사용자가 이미지를 검색하고자 할 때에 그 사용자의 정보와 질의 정보를 데이타베이스에 자동으로 기록하게 하고, 그 기록된 정보로부터 새로운 질의 패턴을 생성할 수 있도록 하는 것이다.

본 논문에서는 이미지 기술 정보를 개념 계층을 이용하여 표현하고 사용자의 질의 정보를 자동으로 추출하여 분석하고 이 결과를 사용자에게 피드백(feed back)함으로서 지능적인 이미지 검색을 수행하는 시스템인 IIRDB(Intelligent Image Retrieval Database)를 설계 구현한다. 본 시스템에서 이미지 검색은 사용자로부터의 지식을 추출하고, 추출된 데이타를 데이타베이스에 저장하며, 지식탐사를 위해 주어진 데이타베이스로 부터 대상이 되는 데이타를 선별하고, 선별된 대상 데이타를 마이닝(mining)하여 구한 결과를 이해하기 쉽게 표현하는 과정으로 이루어진다. 이러한 질의를 지능적으로 사용하기 위한 지능적 사용자 인터페이스도 웹 브라우저를 통해 제공된다.

본 논문의 구성은 다음과 같다. 2장에서는 관련 연구에 대해서 살펴보고, 3장에서는 지능적 이미지 검색 시스템 설계에 따른 요구사항과 혼합형 이미지 메타데이타 모델링 절차를 살펴보고, 이미지 메타데이타 생성과 사용자 메타데이타 생성 방법 그리고 지능적 질의 처리에 대하여 설명한다. 4장에서는 지능적 이미지 검색 시스템 설계에 따른 개발 플렛폼과 지능적 이미지 검색의 예를 보이고, 시스템의 성능평가를 수행한다. 마지막으로, 5장에서는 결론과 향후 연구에 대해 기술한다.

2. 관련 연구

이미지 데이타를 효율적으로 관리하고 사용자에게 제공해야할 필요성 때문에 최근 신속하고 효율적인 이미지 데이타 검색에 관한 연구가 활발히 진행되고 있다. 이러한 연구의 하나로 최근 발전되어가고 있는 지식 기반 혹은 데이타 마이닝 이란 기술을 적용해 좀더 효율적이면서 지능적인 시스템을 만들고자 하는 시도로 이루어지고 있다. 본 절에서는 기존에 연구되었던 이미지 검색 시스템의 특징들과 지능적인 시스템을 만들기 위해 필요한 지식 기반 기술들에 대해 살펴보기로 한다.

2.1 기존의 이미지 검색 기술

이미지 데이타로 부터 추출할 수 있는 특징 데이타들로는 이미지의 모양, 색상, 질감등과 같은 기본적인 특징과 이미지의 다양한 의미 정보(semantic information)를 텍스트 형태로 표현하는 논리적인 특징으로 나눌 수 있다[4]. 기존의 이미지 검색 시스템들은 대부분 내용 기반 검색과 논리적인 특징에 기반한 검색 방법만을 제공하고 있다. 또한, 이미지 처리 기술을 이용하여 이미지 데이타로부터 기본적인 특징을 자동으로 추출하여 이를 기반으로 검색을 수행하는 방법들이 생겨 나고 있으며, 이러한 검색 방법은 이미지 데이타의 특징이 행렬이나 벡터의 형태로 표현되기 때문에 완전일치 검색(exact match)이 아닌 유사 검색(approximate match 또는 similarity retrieval)을 수행하여 사용자가 원하는 이미지와 유사한 이미지들을 검색한다[5]. 이런 내용기반 검색 조차도 사용자의 의도와 만족도와는 상관 없이 검색을 수행하기 때문에 검색 후에도 잘못 찾아진 이미지 때문에 시간을 낭비하거나 검색의 적중률을 감소시키는 문제가 발생하였다. 기존의 내용기반 검색에 대하여 살펴보면 다음과 같다.

QBIC은 IBM Almaden 연구소에서 개발된 내용

기반 이미지 검색 시스템으로, 이미지의 색상이나, 질감, 모양 등과 같은 기본적인 특징에 기반한 다양한 시각적 질의를 제공한다[6,7]. VisualSEEK는 WWW를 이용한 내용 기반 이미지 검색 시스템으로 Columbia 대학에서 개발되었다[8]. 이 시스템의 가장 큰 특징은 시각적 특징(visual feature)에 의한 내용 기반 질의(content-based retrival)와 공간 정보에 의한 공간 질의(spatial retrival)가 동시에 가능하다는 것이다. Jacobs 등이 제안한 내용 기반 이미지 검색 시스템은 특징 데이타의 추출 과정에서 처음으로 웨이브릿 변환을 사용하였다[9]. 이 연구에서는 칼라 이미지(color image)에 대하여 직접 웨이브릿 변환을 적용하여 얻은 계수(coefficient)들을 특징 벡터로 사용하였다.

2.2 지식 탐사 기술과 배경 지식

지식 탐사(knowledge discovery)란 대용량의 단편적인 데이터베이스에서 유용한 정보를 추출하고 이용하는 일련의 과정을 의미하는 용어이다. 지식탐사의 응용 분야로는 정보관리, 질의 처리, 의사결정, 프로세서 제어 등 다양하며 최근 발전하는 온라인 정보 서비스와 웹 정보 서비스에서도 보다 나은 정보의 제공과 사용자의 행동양식(access pattern)을 파악하는 등의 용도로 활용되고 있다[10]. 지식탐사 유형에 대표적인 데이타마이닝 기술은 다음과 같다.

- 연관규칙 (association rule)[11,12]
- 순차 패턴 (sequential pattern)[13]
- 순회 패턴(traversal patterns)[14]
- 분류 패턴(classification pattern)[15]

분류 패턴은 귀납적 학습문제에서 가장 많이 연구가 되어진 분야로서 각 클래스가 갖는 특징에 근거하여 데이타를 분류하는 것이다. 분류화를 표현하기 위한 방법으로는 결정 트리 접근(decision tree approach) 방법이 가장 많이 사용된다. 데이타마이닝의 한 기법인 분류(classification)는 어떤 집합을 여러 개의 서로 다른 부분집합으로 나누어 각각을 어떤 특정한 성질을 가지는 클래스로 구분짓는 것으로, 각 클래스에 속하는 구성원소들은 다른 집합에 존재하는 원소들보다 높은 유사성을 지니게 된다. 분류를 하는 방법으로는 의사결정 트리(decision tree) 방법과 신경망(neural network) 방법, 그리고 유전 알고리즘(genetic algorithm)을 이용한 방법 등이 있다[16].

3. 지능적 이미지 검색 시스템 설계

앞에서 언급한 바와 같이 이미지 데이타는 대 용량적인 특성과 비정형적인 특성을 가지고 있기 때문에 이미지에 대한 표현을 메타데이타로 표현하는 것이 바람직하다. 따라서, 3장에서는 지능적 이미지 검색 시스템의 설계를 위한 요구사항, 이미지 메타데이타를 표현하기 위한 개념 계층 구조, 사용자의 정보와 질의 데이타를 자동 추출하기 위한 구조, 지능적 질의 처리 단계를 설명한다.

3.1 지능적 이미지 검색 시스템 요구사항

본 시스템의 사용자는 특정 분야에 국한되지 않은 사용자로 가정한다. 이미지의 메타데이타와 묘사 정보는 이미지에 대한 영역 전문가나 지식 공학자에 의해 주어진 데이타를 사용한다고 가정한다. 또한, 이미지의 내용 기반 검색을 위한 이미지의 특징 정보는 이미지의 전처리(preprocessing) 과정을 거쳐 이미지 데이타베이스에 저장되어 있는 것으로 가정 하였다. 입력된 원 화상은 임의 크기의 Ture 칼라나 256 칼라의 GIF나 JPEG화일 형식이다.

본 논문의 지능적 이미지 검색 개발 방법은 기존의 내용 기반 이미지 검색과 텍스트 기반 검색을 통합하고, 이미지를 검색하기 위한 지능적 인터페이스를 사용자에게 제공한다. 지식 기반을 기초로한 마이닝 기법이나 분석통계 방법을 이용하여 사용자 정보를 분석하고, 이미지를 검색하였던 사용자의 질의 정보나 에러 정보를 자동 추출할 수 있는 방법을 연구하여, 지능적 이미지 검색 시스템에 적용할 수 있는 기술을 개발한다.

사용자의 만족도나 정확도를 고려하지 않은 이미지 검색을 좀 더 지능화하고 지식 개발 기술을 통합한 검색을 지원하는 시스템을 개발하기 위해 다음과 같은 사항을 고려한다.

- 사용자에게 다양하고 지능적인 이미지 검색 인터페이스를 제공한다.
- 이미지 메타데이타와 설명 데이타를 통합하여 이미지의 검색 효율과 사용자의 만족도를 높인다.
- 사용자에게 다양한 정보를 제공하기 위해 사용자

로부터 질의 정보를 자동으로 추출하여 데이타베이스에 저장할 수 있게 한다.

- 사용자 정보와 이미지 메타데이타를 혼합 이용하여 좀 더 지능적이고 효율적인 검색을 할 수 있게 한다.
- 사용자로부터 얻어진 질의 정보를 다양하게 분석하여 다른 사용자에게 유용한 정보가 되도록 한다.
- 사용자가 잘못 접속하거나 검색하였을 때 시스템이 지능적으로 대처할 수 있도록 한다.
- 사용자의 접속 통계나 질의 정보를 마이닝하여 검색할 수 있도록 하므로써, 일반 사용자에게 데이타 마이닝의 결과나 통계 결과를 피드백시킨다.

이러한 기능을 수행할 수 있도록 본 논문에서는 그림 1과 같은 지능적 이미지 검색 시스템 구조를 제안한다.

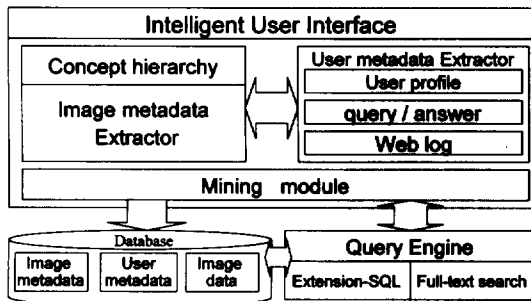



그림 1. 지능적 이미지 검색 시스템 구조

사용자 메타데이타 추출기(user metadata extractor)는 사용자 아이디(ID)나 사용자의 IP, 사용시간, URL등의 정보를 자동으로 추출하여 데이타베이스에 메타데이타(metadata) 형태로 기록한다. 즉, 웹 로그(web log)를 데이타베이스에 기록 저장하는 것이다.

이러한 사용자 로그정보를 수집하는 단계는 사용자가 처음으로 시스템에 접속하였을 때 이루어지며, 사용자의 마지막 요청후 일정시간 후에 종료가 된다. 이미지 메타데이타 추출기(image metadata extractor)는 이미지에 대해서 그 이미지에 관한 정보를 개념 계층(concept hierarchy)에 기반하여 표현되고, 이미지 자체와 그 이미지에 대한 메타데이타를 구분하여 데이타베이스에 저장한다. 새로운 이미지가 등록될 경우에는 이미지 데이타와 이미지 메타데이타를 개념 계층에 의한 분류를 하여 웹상에서 지능적으로 등록하기 위한 지능적 사용자 인터페이스(intelligent user interface)가 제공된다. 마지막으로 사용자의 정보를 이용하여 지능적으로 이미지를 검색하기 위해서 사용자가 질의한 이미지와 사용자 ID, 검색 시간을 데이타베이스에 순차적으로 기록한다. 만일, 다른 사용자가 같은 질의를 하거나 비슷한 질의를 하였을 때, 사용자의 정보와 질의 정보를 다양하게 분석하여 부가적인 정보를 제시한다.

3.2 혼합형 이미지 메타데이타 모델링

지능적 이미지 검색 시스템을 구현하기 위해서 우선 이미지의 데이타를 효율적인 방법으로 데이타베이스에 저장할 수 있어야 한다. 또한, 사용자로부터 자동으로 추출된 사용자 질의 정보는 사용자의 ID와 시간 그리고 질의한 이미지 정보를 데이타베이스에 자동적으로 저장할 수 있어야 한다. 사용자의 메타데이타는 크게 세가지로 구분된다. 즉, 사용자의 신상정보를 나타내는 사용자 정보, 사용자가 이미지를 검색하여 이미지를 선택하였을 때의 기록을 나타내는 사용자 질의 정보, 그리고 사용자가 시스템에 접속하였을 때의 기록을 나타내는 사용자 로그 정보로 구분된다. 이미지 메타데이타는 개념 계층 정보, 이미지에 대한 설명 정보, 그리고 이미지의 특징 정보로 구성된다.

3.2.1 이미지 메타데이타

지식 탐사는 대용량의 데이타를 대상으로 하기 때문에 많은 계산을 필요로 하고, 결과로 얻는 지식의 질(quality)이 중요한데 배경 지식을 사용함으로써 지식을 찾아내는데 필요한 탐색공간을 줄이고 탐색의 결과로 보다 흥미로운 지식을 얻을 수 있게 된다. 배경 지식의 대표적인 예가 개념 계층(concept hierarchy)이다.

개념 계층(CH: concept hierarchy)이란 개념 공간상에 존재하는 개념들을 레벨에 따라 가장 일반적인 개념에서부터 그에 대응되는 구체적인 개념순(general -to-specific ordering)으로 나열한 것이다. 개념 계층은 주로 트리(tree), 격자(lattice), DAG(directed acyclic graph) 등의 형태를 가진다. 개념 계층이 지식탐사에서 대표적인 배경 지식으로 사용되는 이유는 그 형태가 간단하고, 일반화(generalization) 과정에 적합하기 때문이다. 개념 계층은 영역 전문가(do-

main -expert)나 지식 공학자(knowledge engineer)에 의해 주어지는 것이 일반적이나 데이타의 분포를 고려하여 자동적으로 생성되기도 한다[17,18].

그림 2는 이미지 개념 계층의 일반화 관계를 표현하기 위한 데이타베이스 스키마 이다.

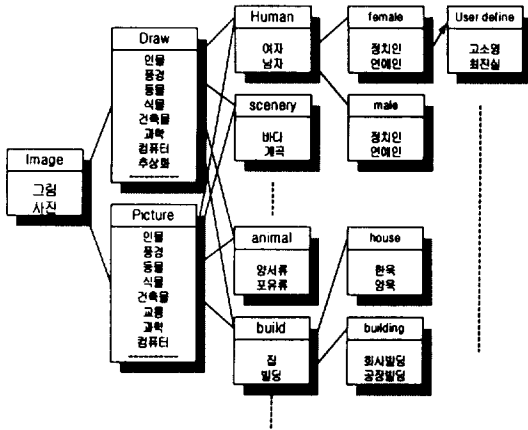



그림 2. 이미지 개념 계층을 위한 스키마

이미지에 대한 개념 계층만으로 이미지의 정보를 표현하기에는 부족하기 때문에 이미지 설명 정보를 추가하고, 내용기반 검색을 통합하기 위해 이미지의 평균 칼라값인 RGB(red, green, blue) 칼라값을 이미지의 히스토그램으로부터 추출하여 데이타베이스에 저장하였다. 그림 3은 SQL sever에 ODBC를 통하여

(a) 이미지 메타데이타의 개념 구조

<Image data>

이미지 ID	등록자	주제	검색횟수	등록일	설명정보	사이즈	파일 이름	이미지 소스
IID	name	subject	visited	date	content	size	filename	isource
int	varchar	varchar	int	date	text	varchar	varchar	image

<Image semantic data>

이미지 ID	대분류	중분류	소분류	사용자정의	설명정보	객체정보	느낌정보	동작정보	장소정보
IID	class1	class2	class3	UD	b_infor	o_infor	I_infor	m_onfor	a_onfor
int	varchar	varchar	varchar	varchar	varchar	varchar	varchar	varchar	varchar

<Image feature data>

이미지 ID	Red	Green	Blue
IID	Rcolor	Gcolor	Bcolor
int	varchar	varchar	varchar

(b) 이미지 메타데이타의 논리적 구조

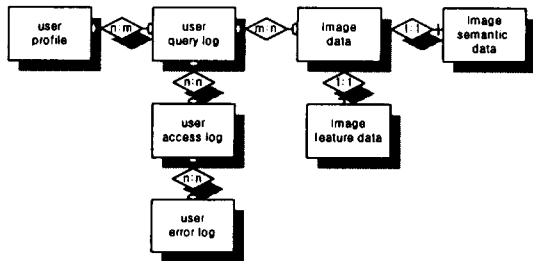
그림 3. 이미지 메타데이타 모델링

이미지의 특징 데이타와 설명 데이타를 저장하기 위한 이미지 메타데이타 모델링 결과를 보이고 있다.

3.2.2 사용자 메타데이타

본 지능적 이미지 검색에서는 이미지의 정보 뿐만이 아니라 사용자의 정보를 반영하여 이미지를 검색하는 기법을 제시하였다. 사용자에게 다양한 질의를 할 수 있게 할 뿐만 아니라, 자신과의 비슷한 환경이나 취미를 갖은 사람들이 찾았던 이미지를 찾게 하므로써, 사용자의 의도를 반영하고 사용자의 질의 정보를 피드백하여 다른 사용자에게 도움을 줄 수 있도록 한다. 또한, 좀더 지능적이고 효율적인 시스템을 구현 하기 위해 사용자의 에러정보를 기록하게 함으로써, 시스템은 관리자에게 이러한 정보를 분석하여 알려, 관리자가 지능적인 대처를 할 수 있도록 한다. 그림 4는 사용자가 회원 등록시 자신의 정보를 등록할 사용자 정보와 사용자가 이미지를 검색하였을 경우 사용자 질의 정보, 사용자가 접속하였을 때 사용자의 로그를 기록하기 위한 사용자 로그 정보, 그리

(a) 사용자 메타데이타의 개념 구조

<user profile>

사용자 ID	암호	성	이름	나이	생년	월	일	성별
UID	passwd	fname	lname	age	b_year	b_month	b_day	sex
varchar	varchar	varchar	varchar	varchar	varchar	varchar	varchar	varchar

결혼여부	국가	주(도)	시	주소	직업	취미	전화	핸드폰	E-mail	등록일
m	nation	state	city	address	job	hobby	phone	h_phone	email	date
varchar	varchar	varchar	varchar	varchar	varchar	varchar	varchar	varchar	varchar	date

<user query log>

사용자 ID	이미지 ID	IP	검색시간
UID	IID	ip	time
varchar	int	varchar	time

<user error log>

ErrorTime	Error URL	Target	HostIP
ErrorDateTime	navError	target	ip
date	int	varchar	varchar

<user access log>

사용자접속 시간	주단위 기록	일별 기록	HostIP
Time	Week	date	ip
time	varchar	date	varchar

(b) 사용자 메타데이타의 논리적 구조

그림 4. 사용자 메타데이타 모델링

고 사용자의 에러를 기록하기 위한 사용자 에러 정보를 저장하기 위한 사용자 메타데이타 모델링 결과를 나타낸다.

3.3 지능적 이미지 검색 시스템

지능적 이미지 검색 시스템 구조는 우선 사용자에게 다양한 질의를 할 수 있는 지능적 이미지 검색 인터페이스인 Intelligent Image finder(IIfinder)를 제공한다. 지능적 사용자 인터페이스는 사용자에게 검색의 폭을 줄여 주면서 사용자가 찾고자 하는 이미지의 유형에 따라 검색할 수 있도록 한다. 이러한 검색 방법을 통합하여 질의할 수 있는 다단계 질의 방법도 제공한다. 사용자로부터 입력된 질의 데이타는 웹을 통해 이미지 검색 시스템으로 전달되고, 이때 사용자의 메타데이타를 추출하기 위한 사용자의 정보 추출기가 데이타베이스에 자동으로 메타데이타 형태로 저장하게 된다. 저장된 메타데이타는 다양한 방법으로 마이닝되어 사용자에게 제공되어지고 사용자로부터 입력된 검색 데이타는 지능적 질의 처리에서 Transact-SQL 형태의 질의어를 생성하고 데이타베이스에 질의하여 질의 결과를 사용자에게 HTML 형식으로 보여주게 된다. 이미지의 데이타는 축소되어 축약이미지(thumbnail) 형식으로 보여지게 된다. 그림 5는 질의 인터페이스와 질의 데이타 추출과정, 질의 처리 과정을 보이고 있다.

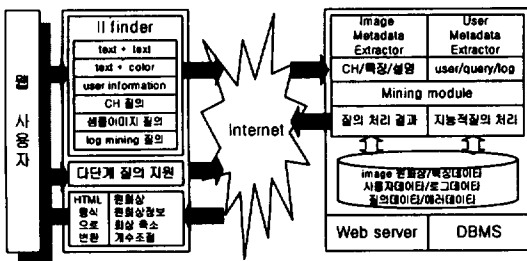



그림 5. 지능적 이미지 검색 시스템의 처리 절차

3.3.1 혼합형 이미지 메타데이타 생성

이미지 메타데이타 생성 절차는 웹을 통한 이미지의 등록을 통해 이미지의 메타데이타를 생성한다. 이미지를 등록하기 위한 절차는 두단계로 볼 수 있다. 이미지 원화상을 시스템으로 업로드하고, 그 이미지의 설명 데이타나 특징 데이타를 등록하게 된다. 이렇게 등록된 데이타는 이미지 데이타베이스에 저장되게 한다. 사용자 메타데이타 생성 절차는 시스템에 접속하여 이미지를 검색하고자 하는 사용자는 시스템에서 회원으로 관리되고, 사용자의 행동 양식을 관리하여 시스템을 지능적이고 효율적으로 만든다. 이를 위하여 사용자의 정보 입력을 필요로 하며, 입력된 정보는 질의 대상 데이타가 된다. 사용자의 질의 패턴에 따른 사용자들의 분류와 클러스터링을 수행함으로써 사용자들에게 부가적으로 관련된 정보를 제공한다. 사용자의 질의 정보를 이용하기 위해 사용자의 질의 정보를 자동으로 추출하여 데이타베이스에 저장하기 위한 과정이 필요한데, 이것은 시스템의 환경 변수를 이용하여 사용자의 IP나 접속 시간대를 자동으로 알아 낼 수 있으며, 사용자의 정보를 추출하기 위해 사용자가 로그온시 사용한 사용자ID를 사용한다. 이러한 사용자 모델링을 통해 사용자의 검색에 대해 마치 사람이 대답하는 것과 같이 좀 더 의미 있는 결과와 부가적인 정보를 제공하는 협조적 응답(cooperative answering) 방식을 제공할 수 있다. 그림 6은 사용자로부터의 메타데이타를 생성하기 위하여 자동으로 사용자 정보를 추출하는 루틴을 나타낸다. 이러한 과정을 거쳐 사용자에 대하여 필요한 정보나 데이타를 데이타베이스에 저장하고, 저장된 데이타는 시스템의 효율적인 관리와 검색 서비스를 위해 사용된다. 사용자는 자신의 ID와 패스워드를 입력하고 이미지 검색을 할 수 있다. 이때 시스템은 데이타베이스로부터 사용자 아이디를 체크하게 되고, 사용자로 인식되면 시스템은 사용자에게 개인 세션(session)을 부여하고 사용자의 로그를 먼저 기록하게 된다. 사용자는 이미지를 검색하고, 검색된 이

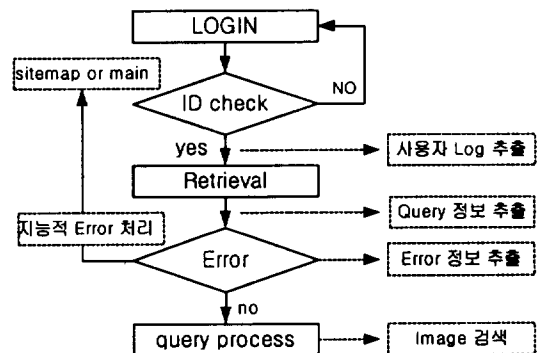



그림 6. 사용자 메타데이타 자동 추출 과정

미지로부터 자신이 원하는 이미지를 선택하게 될 것이다. 이때 시스템은 사용자의 개인 세션으로부터 ID와 이미지의 ID와 시간을 자동으로 데이타베이스에 기록하게 된다.

그림 7은 사용자 질의 정보를 자동으로 추출하여 데이타베이스에 기록하게 되는 루틴을 설명한다. 사용자의 질의 정보는 다른 사용자에게 피드백(feed-back)되어 검색 대상의 데이타가 될 수 있다. 만일, 검색하려는 사용자와 비슷한 환경의 사람이 검색했었던 이미지를 찾고자 할 경우 기존의 이미지 검색 시스템에서는 이러한 검색을 할 수 없었다. 왜냐하면, 사용자의 의도를 반영하지 않았기 때문이다. 본 지능적 이미지 검색 시스템에서는 사용자의 질의 정보를 자동으로 추출하여 다른 사용자에게 피드백 함으로써 다른 사용자의 정보를 이용한 검색을 가능하게 한다. 예를 들면 "20대 남자이고 취미가 영화 감상인 사람이 검색 했었던 이미지를 찾아라?"라고 했을 때 자신과 비슷한 취미와 나이를 가지고 있다면, 아마도 자신이 찾고 싶었던 이미지를 누군가 먼저 찾았었고, 시스템은 그 이미지 정보를 제공하므로써 좀더 지능적인 결과를 제시할 수 있다. 또한, 여러 사용자의 검색 패턴을 알 수 있는데, 만약 취미가 영화감상이고 젊은 20대층의 사람들은 영화 포스터를 취미로 모으거나 자신의 홈페이지에 영화와 관계된 이미지를 삽입하기 위해 영화와 관계된 이미지들을 주로 찾았었음을 알 수 있다.

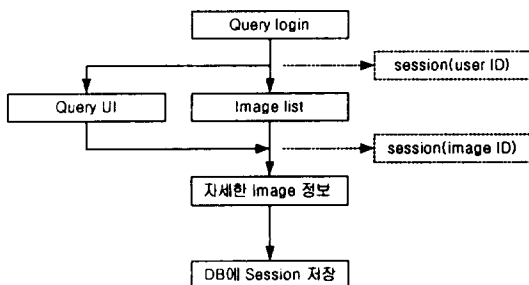
또한, 좀더 효율적이면서 지능적인 이미지 검색 시스템을 만들기 위해서는 우선 사용자가 자주 범하는 오류 발생과 언제 어느 때에 자주 접속하는 경향이 있는가를 파악하는 것이 중요하다. 만일, 사용자가 어느 특정 페이지에 접근하려고 할 때마다 오류가 발생하였다면 시스템 관리자는 새로운 웹 페이지 설계를 해야 한다. 또는, 평소 5시에서 6시 사이에 가장 많은 접속자들이 이미지 검색을 하기 위해 지능적 이미지 검색 시스템에 접속하는 경향이 있다고 한다면, 시스템 관리자는 그 시간대에 가장 좋은 성능을 발휘 할 수 있는 시스템 환경을 만들어 놓아야 할 것이다. 사용자 입장에서는 다른 사용자들이 많이 사용하는 시간대를 피하여 사용자 수가 적은 시간을 이용한다면, 좀 더 빠른 검색과 최대의 시스템 효율로 검색을 할 수 있게 된다. 즉, 대량의 이미지 데이타를 처리해야 하는 만큼 시스템의 성능 저하를 유발 할 수 있는 시스템 사용자 수의 집중에 지능적으로 대처할 수 있다. 자동으로 추출된 사용자 정보나 질의 정보 그리고 이미지 데이타들은 특정 열(column)이나 행(row)으로부터 검색이 가능하도록 해야 하며 이렇게 검색된 정보는 다시 특정 조건에 따라 그룹 지어 지거나, 연산을 수행하여 사용자에게 전달되어 진다.

그림 7. 사용자 질의 로그 데이타 자동 추출 과정

3.3.2 지능적 질의 처리

지능적 질의 처리란 시스템이 사용자의 질의에 대해 구문적인 질의의 답만을 제공하는 것이 아니라 좀더 많은 정보 또는 사용자의 의도를 추측하여 사용자의 의도를 반영한 정보를 제공한다고 말할 수 있다. 또한, 지능적 질의 처리는 사용자에게 질의 내의 스키마 오류에 대한 영향을 줄여주고 사용자가 의도했던 질의의 결과를 유추하여 제공해야 한다. 지식 탐사 기술은 사용자들이 자주 범하는 스키마 오류에 대해 요약된 정보를 생성하여 관리자에게 제공하여 시스템 관리에 필요한 정보를 제공 할 수 있게 한다. 이러한 지능적 질의 처리를 하기 위해서는 시스템에 사용자가 사용하고 있는 정보를 추적하고, 그 결과를 데이타베이스에 저장하므로써 사용자의 질의에 대한 지능적 질의 처리를 수행 할 수 있다. 시스템에서 자주 발생되는 에러는 사용자들이 되반복 하여 접속하려는 시도나 검색하려는 경향이 있을 것으로 추정할 수 있으므로, 시스템 관리자는 이에 대한 시스템 재 설정이 필요할 것이다. 사용자로부터 입력된 검색 데이타는 검색 유형에 따라 질의 처리 프로시저에 의해 질의문이 생성되고 처리된다.

본 논문에서 제안한 IIRDB는 다양한 형태의 이미지 검색 기능을 제공한다.

- **T+C 검색(텍스트와 칼라 혼합)** : 이미지에 대한

설명 정보나 속성 검색을 하고자 하는 사용자를 위해 이미지에 대한 통합 검색 방법을 지원한다. 기존의 텍스트 기반과 내용기반 검색이 분리되어 검색하던 방식을 통합함으로써 검색의 정확도를 높이고, 사용자의 만족도를 높일 수 있다. 그림 8은 이미지 설명 정보와 속성 정보를 통합한 인터페이스와 검색 루틴을 나타낸다.

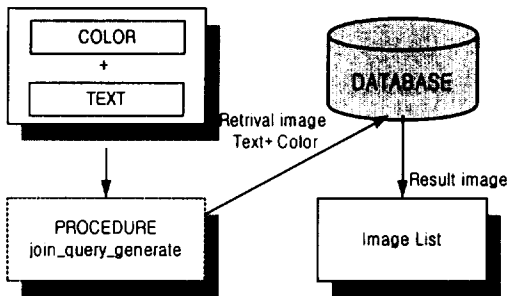



그림 8. 통합 검색 질의 루틴

이미지의 속성과 설명 데이타로부터 통합 검색을 하려면 이미지 메타데이타의 데이타베이스로부터 질의문을 조인하여 질의하여야 한다. 이러한 처리를 하기 위한 프로시저는 다음과 같다.

```
Procedure Join_Query_Generate
  begin
     database connection.open;
     사용자로부터 질의 데이타 추출 ;
     join T-SQL문 생성;
         if (rs.EOF = true) then
              response.redirect
         else
                 do while not rs.EOF
                      rowcount=rowcount+1
                      rs.movenext
                 loop
         database connetion.close
         set rs.nothing
         end if
end  Join_Query_Generate
```

• **UM+IM 검색(사용자 메타데이타와 이미지 메타데이타 혼합 검색)** : 사용자 질의 검색으로 이미지 메타데이타와 사용자 메타데이타를 이용하여 지능적 질의 처리에 의한 검색 방법을 사용자에게 제공한다. 이미지를 검색 했었던 사용자의 취미가 무엇이고, 직업이 무엇인가? 그리고 언제 어떤 이미지를 검색 했었는가? 하는 것을 검색 할 수 있다. 그림 9는 사용자 메타데이타를 이용하여 검색 할 수 있는 인터페이스와 루틴을 나타낸다. 이러한 검색 기법은 기존에 이미지 검색에서는 고려 되지 않은 검색 방법이므로 많은 연구가 계속 되어질 필요가 있다고 본다.

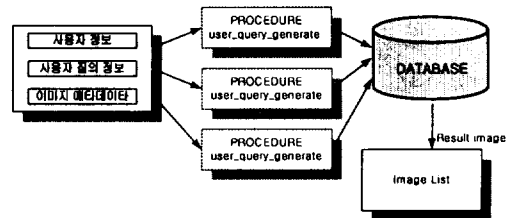



그림 9. 혼합 메타데이타를 이용한 검색 질의 루틴

또한, 사용자의 검색 패턴이나 나이별, 혹은 직업별로 분류를 할 수 있기 때문에 별도의 마이닝 알고리즘이나 연산을 필요로 하지 않는다.

사용자의 메타데이타로부터 검색하기 위한 프로시저는 다음과 같다.

```
Procedure User_Query_Generate
  begin
   database connection.open;
   사용자로 부터 질의 데이타 추출 ;
   이미지 메타데이타 와 사용자 메타데이타로 부
터의 질의문 생성;
   join T-SQL문 생성;
     if (rs.EOF = true) then  response.redirect
       else
         do while not rs.EOF
            prune step      //중복데이타 제거
            rowcount=rowcount+1
            rs.movenext
         loop
       end if
   database connetion.close
   set rs.nothing
 end User_Query_Generate
```

• **CH 검색(개념 계층 이용 검색)** : 이미지를 개념 계층에 의한 클래스로 검색을 할 수 있다. 막연히

특정한 주제가 없이 이미지를 검색하려고 하는 사용자에게 이미지 검색의 혼동을 막고, 개념계층에 의한 분류 방식에 의해 검색 할 수 있기 때문에 쉽게 원하는 이미지를 얻을 수 있을뿐만 아니라, 검색의 대상폭이 적어지므로 효율적인 이미지 검색을 할 수 있다. 예로, 사진이라는 클래스를 지정하였다면, 검색 범위는 사진에 해당되는 이미지만 선택 될 것이고, 사진들 중에도 동물 클래스에 관한 사진이라고 선택한다면 검색 범위는 동물 사진만 검색 대상으로 하기 때문에 검색의 대상이 대폭 줄어 들 것이다. 상위 클래스에 선택된 것에 따라서 하위 클래스가 지능적으로 지정되기 때문에 사용자에게 찾고자 하는 이미지를 쉽게 제공할 수 있다. 그림 10은 개념 계층에 의한 이미지 검색을 제공하는 지능적 인터페이스 와 루틴을 나타낸다.

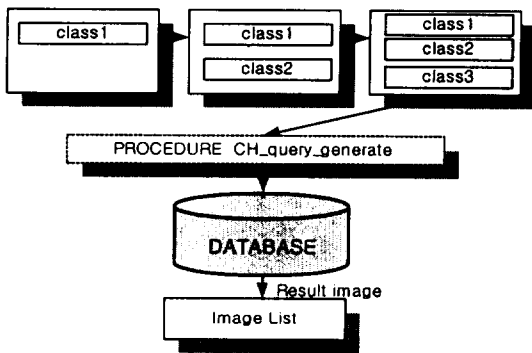



그림 10. 개념 계층에 의한 검색 질의 루틴

개념 계층에 의한 검색 루틴에서 질의 처리를 위한 프로시저는 다음과 같다.

```
Procedure CH_Query_Generate
 begin
  Database Connection.Open;
  사용자로 부터 Class1 data 추출 질의문 생성 ;
   begin
     do while not rs.EOF
          rs("item")
          rs.movenext
     loop
  end
  begin
  if item="그림" then 그림table로부터data 추출;
         do while not rs.EOF
             rs("item")
             rs.movenext
         loop
   else if item="사진" then 사진table로부터data 추출;
         do while not rs.EOF
             rs("item")
             rs.movenext
         loop
    end if
             session(class1)에 class1 선택값 저장 ;
             session(oper)에 and, or 연산자 저장;
             <이미지 클래스 반복 루틴>...
             session() 데이타에 의한 질의문 생성 ;
             database connection.close
             set rs.nothing
         end
   end CH_Query_Generate
```

- **SI 검색(샘플 이미지에 의한 검색)** : 찾고자 하는 이미지와 유사한 형태나 칼라를 직접 선택하고, 선택되어진 이미지의 정보를 데이타베이스로부터 추출하여 검색 데이타로 이용함으로써, 유사한 이미지들을 찾을 수 있다. 그림 11은 샘플 이미지에 의한 검색 방법을 나타내는 인터페이스와 루틴을 나타낸다. 샘플 이미지에 의한 검색 처리 절차는 샘플 이미지중에 선택된 이미지의 데이타를 데이타베이스에 먼저 질의하여 데이타를 생성하고, 생성된 데이타로 다시 재 정의된 질의문을 생성하여 질의할 수 있게 한다. 이러한 질의 처리 절차에서 칼라 값이나 이미지의 사이즈는 데이타의 범위 값을 근사값으로 재조정된다.

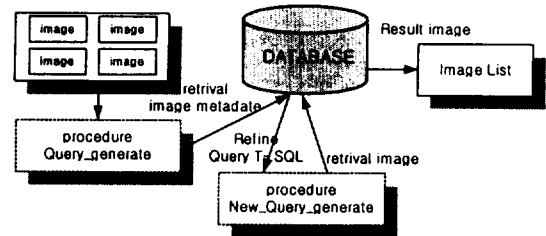



그림 11. 샘플 이미지에 의한 검색 질의 루틴

다음은 샘플 이미지에 의한 질의 처리 절차를 나타내는 프로시저이다.

```
Procedure New_Query_Generate
  begin
    database connection.open;
    선택된 이미지의 정보 추출 질의문;
    set 변수에 선택된 이미지 메타데이타 저장;
    // 값 범위 재 설정
        rcolor1= Cint(rs("rcolor"))-a
        rcolor2= Cint(rs("rcolor"))+a
        gcolor1= Cint(rs("gcolor"))-a
        gcolor2= Cint(rs("gcolor"))+a
        bcolor1= Cint(rs("bcolor"))-a
        bcolor2= Cint(rs("bcolor"))+a
        size1= CLng(rs("size"))-b
        size2= CLng(rs("size"))+b
    close.recordset
  begin
    database.open;
    변수에 의한 이미지 메타데이타 추출 질의문 생성;
    if (rs.EOF = true) then  response.redirect
      else
        do while not rs.EOF
            rowcount=rowcount+1
            rs.movenext
        loop
      end if
          database connetion.close
          set rs.nothing
    end
  end New_Query_Generate
```

4. 시스템 구현 및 평가

앞에서 설계한 지능적 이미지 검색 시스템 설계를 구현 하기위한 전체적인 구조와 기술적인 방법 그리고 이미지 검색의 결과를 제시하고, 그 결과를 성능 평가하여 비교 분석하였다.

그림 12는 지능적 이미지 검색 시스템의 구현 구조를 나타낸다. 사용자는 웹상에서 이미지 검색을 하기 위해 시스템에 등록을 하고, 등록된 사용자의 메타데이타는 데이타베이스에 저장되어 관리된다. 등록된 사용자는 로그인 인증 과정을 거쳐 이미지를 검색하거나, 그 이외의 부가적인 서비스를 받을 수 있다. 이때 사용자의 접속 통계 분석을 하기 위한 사용자 로그를 데이타베이스에 자동으로 기록하게 되며, 사용자는 다단계 검색 인터페이스를 통해 원하는 이미지를 검색한다. 사용자로부터 질의된 이미지는

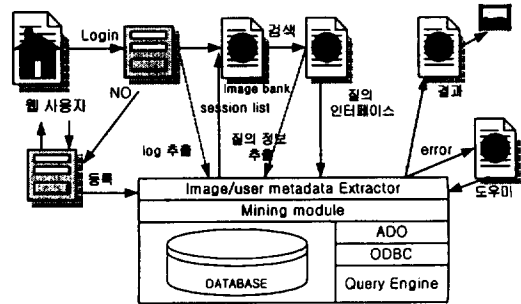



그림 12. 지능적 이미지 검색 시스템의 구현 구조

축약형 이미지와 이미지의 메타데이타와 함께 제공이 되며, 제공된 이미지로부터 선택된 이미지는 사용자로부터 검색 된 이미지로볼 수 있기 때문에 선택된 이미지의 아이디와 사용자 아이디, 그리고 시간을 데이타베이스에 자동으로 저장하게 된다.

이렇게 추출된 데이타는 다른 사용자의 검색 대상이 되며, 사용자의 분석 데이타가 된다. 사용자로부터 입력받은 질의는 ODBC를 통해 Transact-SQL 구문으로 변형되어 질의 엔진으로 전해지게 된다. 질의 엔진은 데이타베이스로부터 각각의 정보를 통합 질의 하여 사용자에게 제공 되어지게 한다.

사용자로부터 질의를 입력받고 입력된 데이타를 데이타베이스에 질의 하기 위해서는 애플리케이션과 데이타베이스를 연결하는 방법이 필요한데, 본 시스템의 구현에서는 이러한 것을 ADO(ActiveX Data Object)객체를 이용하였다[19].

4.1 지능적 이미지 검색의 예

구현된 지능적 이미지 검색 인터페이스에서 제공되는 검색 유형은 여러 가지 형태로 텍스트를 기반으로 하여 이미지 검색이 가능하며, 이미지의 속성과 키워드를 혼합한 통합 이미지 검색, 사용자 메타데이타를 이용한 이미지 검색, 개념 계층에 의한 검색, 샘플 이미지에 의한 검색으로 이루어져 있다. 검색 되어진 이미지는 부가정보와 함께 리스트 형식으로 보여지게 되고, 자세한 이미지의 정보는 해당 주제나 축소형 이미지를 클릭 하여볼 수 있다.

- 키워드 이미지 검색은 주제나, 관련 있는 키워드에 의해 검색한다. 사용자는 키워드와 이미지 등록 날짜 그리고 이미지 크기를 조합하여 AND나

OR와 같은 블리언(boolean)연산을 통해 검색할 수 있다. 통합 이미지 검색은 이미지에 대한 텍스트와 내용 기반 검색을 통합하여 검색한다. 사용자로부터 칼라를 직접 선택하고, 그외의 이미지 설명 정보에 대하여 텍스트로 입력하게 한다. 그림 13은 통합 이미지 검색인터페이스와 검색된 이미지 결과를 나타낸다.

(a) 통합 이미지 검색인터페이스

(b) 통합 이미지 검색결과

그림 13. 통합 이미지 검색 결과 (T+C 검색)

- 사용자 메타데이타 검색은 사용자의 질의 정보와 사용자 정보를 이용하여 이미지를 검색한다. 자신과의 환경이 비슷한 사람이나 다른 사람이 찾았었던 이미지의 정보를 이용하여 검색하고, 사용자 질의 정보를 마이닝하여 질의 할 수 있다. 그림 14는 사용자 메타데이타 검색 인터페이스와 검색된 결과를 나타낸다.

- 개념 계층 이미지 검색은 이미지의 표현을 개념 계층을 이용하여 표현한 데이타베이스로부터 검색한다. 사용자의 상위 클래스 선택에 따라 지능

(a) 사용자 메타데이타에의한 검색 인터페이스

(b) 사용자 메타데이타의한 검색 결과

그림 14. 사용자 메타데이타에 의한 검색

적으로 하위 클래스를 표현해 준다. 그림 15는 개념 계층 이미지 검색 인터페이스와 검색 결과를 나타낸다.

- 샘플 이미지 검색은 사용자가 직접 이미지들을 보고 선택하면 데이타베이스로부터 선택된 이미지의 정보를 추출하여, 추출된 정보로부터 유사한 이미지들을 검색한다. 그림 16은 샘플 이미지 검

(a) 개념 계층 이미지 검색 인터페이스

(b) 개념 계층 이미지 검색 결과

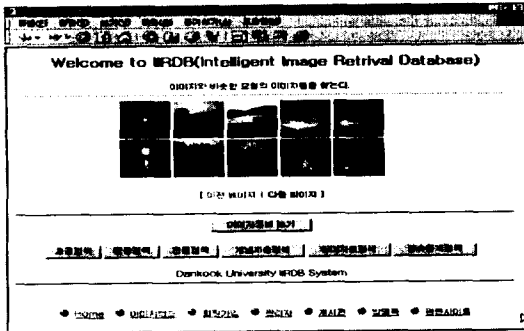
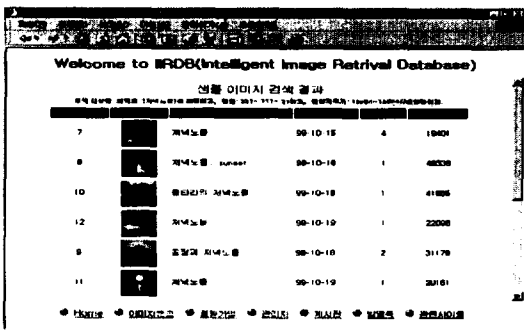
그림 15. 개념 계층 이미지 검색 (CH검색)

(a) 샘플 이미지 검색 인터페이스

(b) 샘플 이미지 검색 결과

그림 16. 샘플 이미지 검색 (SI 검색)

색 인터페이스와 검색 결과를 나타낸다.

4.2 성능 평가

본 지능적 이미지 검색 시스템에서 성능평가를 위해 사용되어진 이미지는 약 200여 개이며 Windows NT환경에 128M 메인메모리의 환경에서 평가하였다. 지능적 이미지 검색의 성능 측정을 하기 위해 검색 시스템의 리콜(recall)과 정확도(precision)측정 하였다.

성능 평가를 하기 위해 그림 17과 같이 사용된 질의는 "저녁노을에 관한 이미지를 찾아라"이며, 데이타베이스에 저장되어 있는 이미지중 관련된 이미지는 7개 이다. 질의중 1과 2는 키워드만을 이용하여 질의(T)하였을 때 결과이며, 3과 4에서는 칼라 속성만을 이용하여 질의(C) 하였을때 결과이고, 5와 6은 키워드와 칼라 속성을 통합(T+C)하여 질의 하였을 때 결과이다. 칼라의 속성은 R, G, B의 평균값을 이용하였기 때문에 ±0.5%와 ±1.5%의 근사값을 적용하여 평가 하였다. 키워드만을 이용하였을때에는 평균적으로 만족도가 약 50%에 가까운 결과를 볼 수 있었으며, 칼라 속성 만을 가지고 검색 하였을경우는 만족도가 비교적 많이 저하 되는 것을 볼 수 있었다. 통합된 검색 방법에서는 키워드와 칼라의 속성 값을 같이 적용하여 평가 하였는데, 만족도가 약 80%에 가까울 만큼 좋아진 것을 볼 수 있었다. 물론, 이미지의 데이타가 계속해서 증가하게 되면, 결과의 수치도 변하게 될 것이다. 그림 18은 검색 실험 결과를 그래프 형식으로 보이고 있다.

[질의] "저녁노을 이미지를 찾아라." (관련이미지 수 : 7개)

no	keyword	color	검색된이미지수	정확히 검색된 이미지 수	recall(%)	precision(%)
1(T1)	"노을"	-	12	5	71.4	41.7
2(T2)	"저녁노을"	-	8	4	57.1	50.0
3(C1)	-	R:224,G:98,B:21 (오차 0.5% 허용)	3	1	14.2	33.3
4(C2)	-	R:224,G:98,B:21 (오차 1.5% 허용)	25	7	100.0	36.0
5 (T+C)1	"저녁노을"	R:224,G:98,B:21 (오차 0.5% 허용)	8	7	100.0	87.5
6 (T+C)2	"저녁노을"	R:224,G:98,B:21 (오차 1.5% 허용)	10	7	100.0	70.0

그림 17. T, C, T+C 검색 실험 결과

그림 18에서 이미지 검색 결과로 알 수 있듯이 키워드와 칼라 속성을 통합하여 질의하였을 때 만족도가 가장 높은 것을 알 수 있다. 이런 통합 검색을 샘플이미지 검색 방법(SI)에 적용하여 그림 19와 같은 리콜과 정확도를 실험하였다. SI1, SI2, SI3는 같은 "저녁노을"이란 주제를 포함하고 있으면서 칼라의 근사값을 ±1.0%, ±2.0%, ±3.0%의 근사값을 각각 적용하였을때의 검색 실험 결과이며, SI4, SI5, SI6는 비슷한 칼라의 값이 변화된 이미지를 같은 근사값으로 각각 적용하였을 때의 실험 결과이다. 관계된 키워드와 함께 칼라의 근사값이 ±2.0%에 가까운 질의에

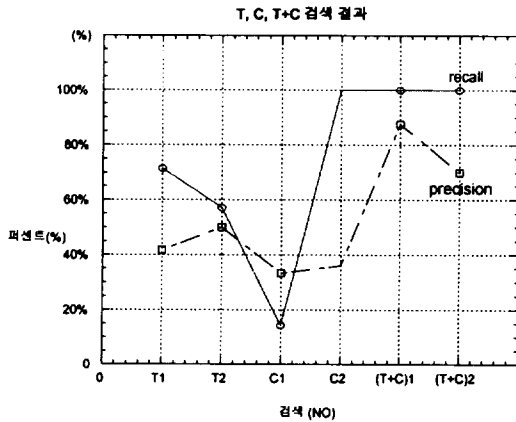
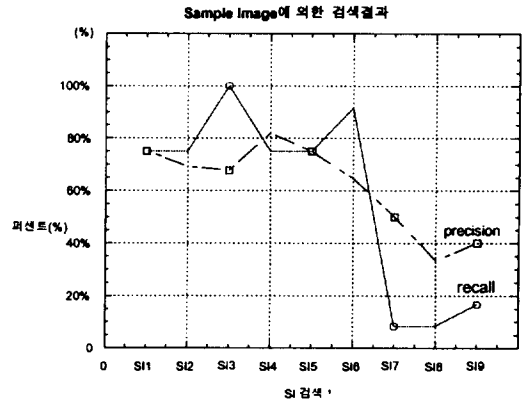



그림 18. T, C, T+C검색 실험 결과 그래프

[질의] "저녁노을과 비슷한 이미지를 찾아라."(관련이미지 수 : 12개)

no	keyword	color	검색된이미지수	정확히 검색된 이미지 수	recall(%)	precision(%)
SI1	"저녁노을"	R:154,G:74,B:16 (오차 1.0% 허용)	12	9	75.0	75.0
SI2	"저녁노을"	R:154,G:74,B:16 (오차 2.0% 허용)	13	9	75.0	69.2
SI3	"저녁노을"	R:154,G:74,B:16 (오차 3.0% 허용)	18	12	100.0	67.7
SI4	"저녁노을"	R:196,G:116,B:47 (오차 1.0%) 허용)	11	9	75.0	81.8
SI5	"저녁노을"	R:196,G:116,B:47 (오차 2.0%) 허용)	12	9	75.0	75.0
SI6	"저녁노을"	R:196,G:116,B:47 (오차 3.0%) 허용)	17	11	91.6	64.7
SI7	"공장과저녁노을"	R:114,G:72,B:41 (오차 1.0%) 허용)	2	1	08.3	50.0
SI8	"공장과저녁노을"	R:114,G:72,B:41 (오차 2.0%) 허용)	3	1	08.3	33.3
SI9	"공장과저녁노을"	R:114,G:72,B:41 (오차 3.0%) 허용)	5	2	16.6	40.0

그림 19. SI 검색 실험 결과

대하여 검색의 만족도와 효율이 높은 것을 알 수 있었다. 반면에 SI7, SI8, SI9의 경우 키워드에 "저녁노을" 이외의 키워드가 포함되어 있을 때 검색의 리콜과 정확도가 저하되는 현상을 볼 수 있다. 검색 결과 이미지에 대한 키워드에 적용이 검색에 큰 영향을 미치고 있다는 사실을 본 실험결과로부터 알 수 있다. 그림 20은 샘플 이미지 검색 방법에 대한 실험 결과를 그래프로 나타내고 있다.

앞의 실험 결과로 다양한 검색 방법으로 사용자에게 질의를 할 수 있도록 하는 검색 인터페이스 제공은 사용자의 만족도를 높일 수 있을 뿐만 아니라, 사용자에게 제공되는 검색의 효율성을 증가시킬 수 있다. 본 논문에서 제시된 지능적 이미지 검색 시스템 IIRDB는 기존의 텍스트나 칼라 기반 검색을 지원하는 동시에 사용자 메타데이타를 혼합하여 검색하는 방식을 제공하며 이미지를 개념 계층에 기반하여 지능적으로 검색하고 있다. 또한, 사용자의 검색 패턴에 따른 효율적인 이미지 검색 시스템을 사용하기 위해 사용자 로그 마이닝 검색을 지원하고 있다.

그림 20. SI 검색 실험 결과 그래프

5. 결 론

혼합형 이미지 메타데이타를 이용한 지능적 이미지 검색은 전자상거래나 이미지 검색 엔진, 전자 미술관 등의 다양한 응용분야에서 사용될 것으로 보인다. 앞으로 웹상에서 이미지 데이타는 계속해서 급증할 것이 확실하며, 이러한 이미지를 효율적으로 관리하고 검색할 수 있는 방법에 대한 연구는 계속될 것이다. 본 논문에서는 대량의 이미지를 데이타베이스에 효율적으로 저장, 관리하고, 사용자 메타데이타와 이미지 메타데이타를 혼합하여 다양한 검색 방법을 지원하면서 지능적으로 검색할 수 있는 IIRDB 시스템을 설계, 구현하였다.

검색 시스템의 성능 측정은 키워드만으로 검색하였을 때, 키워드와 내용기반 검색을 통합한 검색을 하였을 때, 사용자 메타데이타를 이용하여 검색 하였을 때, 샘플 이미지를 이용하여 검색 하였을때의 리콜과 정확도를 비교하여 검색 시스템의 성능 측정을 비교하였다. 성능 측정 결과 키워드와 내용기반 검색을 통합한 검색이 정확도가 가장 높았으며, 샘플 이미지에 의한 검색이 다음으로 정확도가 높았다. 사용자 메타데이타에 의한 검색은 정확도는 떨어졌지만, 사용자에게 유용한 검색방법을 제공하며, 향후 검색 패턴이나 지식탐사에 적용할 수 있는 방법이라는 결론을 얻었다.

본 논문에서 구현된 이미지 검색 시스템은 특정 분야의 이미지 외에, 웹 상에서 사용되는 사진, 그림, 애니메이션등 다양한 이미지를 관리 할 수 있고 이러한 대량의 이미지로부터 필요한 이미지를 지능적으

로 신속하고 정확하게 찾아 줄 수 있는 II finder를 구현 하였다. 기존의 텍스트 검색과 칼라 기반 검색을 동시에 지원하면서 사용자 메타데이타와 이미지 메타데이타를 이용한 검색 방법과 개념 계층에 의한 검색을 지원하고 있다. 또한, 시스템의 효율적인 사용을 위해 사용자 로그 마이닝 검색을 지원하고 있다.

향후 연구과제로 제안시스템을 기반으로 자동으로 추출된 메타데이타를 좀 더 효용성 있게 분석할 수 있는 알고리즘이나 분석방법이 연구되어야 한다. 또한, 다양한 분석과 좀 더 유용성이 있는 검색을 지원하기 위해서는 이미지 데이타를 자동으로 분석 처리하여 유용한 데이타를 추출할 수 있는 알고리즘이나 방법이 필요하다. 만일, 이미지의 특징 데이타들이 자동으로 추출되어 데이타베이스에 저장된다면, 이미지의 전처리를 위한 소요시간을 줄일 수 있을 것이다.

마지막으로 이미지 정보를 기술하는 방법으로 MPEG-7기술을 이용하여 내용 자체 또는 사용자가 관심 있는 자료를 빠르고 효율적으로 검색할 수 있도록 표준화 기법의 연구도 필요하다. 사용자로부터 웹상에 등록되는 이미지 데이타들은 서로 각기 틀린 사용자의 견해나, 사고방식의 차이로 인해 기술 방법이 틀리므로 검색의 효율성을 감소시키는 요인이 된다.

참 고 문 헌

[1] http://drogo.cselt.it/mpeg/standards/mpeg-7/mpeg-7.htm

[2] 홍성용, 나연묵, "웹 환경하에서 지능형 이미지 정보 검색 시스템의 설계 및 구현", 한국멀티미디어학회, 봄 학술발표 논문집, 1999, pp.245-247.

[3] 홍성용, 나연묵, "개념 계층과 사용자 정보를 이용한 지능적 이미지 검색 시스템 설계", 한국정보과학회, 가을 학술발표 논문집, 1999, pp.290-292.

[4] 방난효, 엄기현, "내용 기반 검색을 위한 이미지 데이타베이스 모델", 한국정보과학회, 봄 학술발표 논문집, 1997, pp.11-14.

[5] V. N. Gudivada and V. V. Raghavan, "Content-Based Image Retrieval System," *IEEE Computer*, 28(9), 1995.

[6] W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Petkovic, P. Yanker and C. Faloutsos, "The QBIC Project : Querying images by content using color, texture, and shape," *in Proc. SPIE Storage and Retrival for Image and Video Database,* February 1993, pp.173-187.(http://libra.ucdavis.edu/cgi-bin/QbicStable/)

[7] C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W. Niblack, D. Petkovic, and W. Equiz, "Efficient and Effective Querying by Image Content," *Journal of Intelligent Information System (JIIS)*, July, 1994, 3(3) : 231-262.

[8] J. R. Smith and S.-F. Chang, "VisualSEEk: A Fully Automated Content-Based Image Query System," *ACM Multimedia 96,* Boston, MA, 1996. (http://www.ctr.columbia.edu/)

[9] C. E. Jacobs, A. Finkelstein, and D. H. Salesin, "Fast Multiresolution Image Query," *in Proc. ACM SIGGRAPH,* New york, 1995.

[10] 김성민, 이동하, 남도원, 이전영, "관계형 데이타베이스에서의 지식탐사를 위한 개념 계층의 제어기법", '98 한국전문가시스템 추계학술대회 논문집, 1998년 12월, pp.83-90.

[11] R. Agrawal, R. Srikant, "Fast Algorithms for Mining Association Rules," *in Proc. of the 20th Int'l Conference on Very Large Databases,* Santiago, Chile, Sept. 1994.

[12] J.S. Park, M.-S. Chen and P.S. Yu, "An effective hash-based algorithm for mining association rules," *in Proc of ACM SIGMOD Conference on Management of Data,* San Jose, California, May, 1995, pp.175-186.

[13] R. Agrawal, R. Srikant, "Mining Sequential Patterns," *in Proc. of the Int'l Conference on Data Engineering (ICDE),* Taipei, Taiwan, March, 1995.

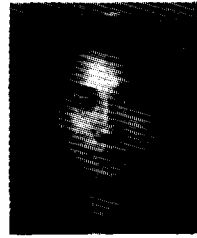
[14] M.-S. Chen, J.S. Park and P.S. Yu, "Data Mining for path Traversal patterns in a Web Environment," *in Proc of the 16th International Conference on Distributed Computing Systems,* Hong Kong, May, 1996, pp.385-392.

[15] 박종수, 유원경, 홍기형, "연관 규칙 탐사와 그 응용", 한국정보과학회 논문지, 제16권 제9호. 1998. 9, pp.37-44.

[16] 김정자, 이도헌, "데이타 마이닝 기술 및 연구동향", 한국정보과학회 논문지, 제16권 제9호, 1998. 9, pp.6-14.

[17] J. Han and Y. Fu, "Dynamic Generation and Refinement of concept Hierarchies for Knowledge Discovery in Database," *AAAI'94 Workshop on Knowledge in Databases (KDD'94)*, Seattle, WA, July1994, pp.157-168.

[18] J. Han, Y. Huang, N. Cercone and Y. Fu, "Intelligent Query Answering by Knowledge Discovery Techniques," *IEEE Transactions on Knowledge and Data Engineering*, 8(3) : 373-390, 1996.

[19] Dina Fleet, Matt Warren, Joyce Chia-Hsun Chen, Alexander Stojanovic, *Active Web Database Programming*, Sams.net, Inc., 1997.

홍 성 용

1998년 우송대학교 컴퓨터과학과 (공학사)
2000년 단국대학교 대학원 컴퓨터공학과 (공학석사)
2000년~현재 단국대학교 컴퓨터공학과 (박사과정)
관심분야 : 멀티미디어 데이타베이스, 멀티미디어 정보 검색, WEB-DB, 데이타 모델링, 전자상거래, 데이타 마이닝

나 연 묵

1986년 서울대학교 컴퓨터공학과 (공학사)
1988년 서울대학교 대학원 컴퓨터공학과 (공학석사)
1993년 서울대학교 대학원 컴퓨터공학과 (공학박사)
1993년~현재 단국대학교 컴퓨터공학과 부교수
1994년~1999년 DASFAA Steering Committee Secretary
1996년~현재 한국정보과학회 데이타베이스연구회 운영위원
1997년~1999년 한국멀티미디어학회 논문지 편집위원
1997년~현재 서울특별시 정보화추진위원회 위원
1999년~현재 한국정보과학회 논문지 편집위원
관심분야 : 데이타베이스, 객체지향 데이타베이스, 데이타 모델링, 데이타베이스 설계, 멀티미디어 시스템, 멀티미디어 데이터베이스, 멀티미디어 정보 검색